

BAB III METODE KERJA

3.1 Waktu dan Tempat

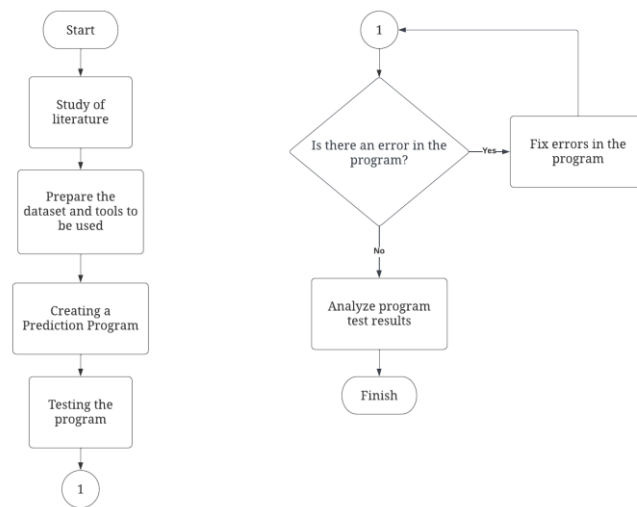
Waktu pengerjaan dilakukan selama 2 minggu dari tanggal 18 April 2022 hingga 29 April 2022.

3.2 Alat dan Bahan

Tools yang digunakan dibedakan menjadi dua yaitu *software* dan *hardware*. Untuk *software* menggunakan google colab yang berfungsi untuk analisa dan visualisasi data dengan bahasa pemrograman *Python*. Sedangkan *hardware* yang dipakai adalah Laptop Sony Vaio dengan spesifikasi layar 14 inch, intel core i3 3217U @ 1.80GHz, memory RAM 8 GB, dan SSD 128 GB.

3.3 Metode dan Proses Kerja

Tahap awal dalam penelitian ini adalah studi literatur dan mempersiapkan dataset, metode, dan *tools* yang akan digunakan. Studi literatur dilakukan untuk mengumpulkan informasi berupa sumber yang kredibel terkait objek yang akan diteliti. Tujuannya agar penulis lebih memahami topik penelitian. Tahap kedua yaitu menciptakan program prediksi dan menguji program. Setelah program berhasil diciptakan maka akan diuji untuk mendapatkan hasil yang mendekati data aslinya. Apabila terdapat *error* seperti *error syntax* maka program akan diperbaiki dan diuji hingga tidak terdapat *error* dan mendapat hasil yang lebih baik. Dan apabila program yang diuji berhasil maka akan di analisis ditahap berikutnya. Terakhir adalah menganalisis hasil pengujian program. Berikut adalah diagram alur penelitian:



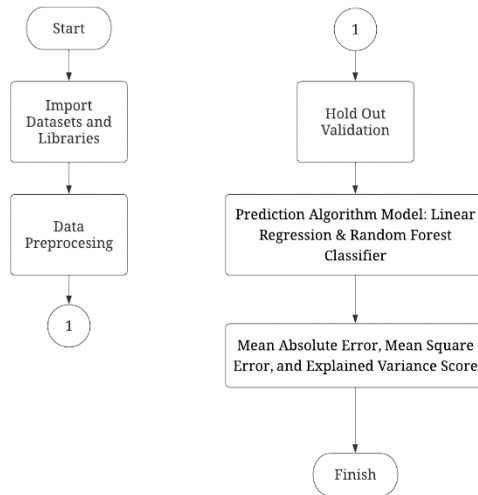
Gambar 3.1 Alur Penelitian

Selanjutnya untuk perancangan program dan metode yang akan dipakai adalah sebagai berikut :

1. *Import Dataset dan Pustaka*. Proses ini bertujuan untuk memasukkan dataset [Uber and Lyft Dataset Boston, MA | Kaggle](#) [1] yang berupa *file* csv dan pustaka yang dibutuhkan untuk proses analisis.
2. *Data Preprocessing* adalah proses awal mengubah data *inputan* menjadi data dengan format yang pas dan siap untuk diproses ke tahap berikutnya. Hal-hal yang dilakukan ketika *preprocessing* ialah: penggabungan, perubahan bentuk, ataupun pentransformasian data sebagai cara untuk membersihkan, mengintegrasikan, mereduksi dan mendiskritisasi. Tujuan proses ini adalah untuk meningkatkan performansi dari model klasifikasi.
3. *Train Test Split/Hold Out Validation*. Pada proses ini dataset [Uber and Lyft Dataset Boston, MA | Kaggle](#) [1] akan dibagi menjadi dua yaitu data *training* yang berjumlah 80% dari keseluruhan data data *testing* yang berjumlah 20% dari keseluruhan data.
4. *Prediction Algorithm Model*. Dalam penelitian ini menggunakan dua algoritma yaitu *linear regression* dan *random forest*. Kedua algoritma tersebut masing - masing akan mengolah data yang didapatkan dari keluaran *train test split*. Hasil akhir dari *linear regression* dan *random forest* akan

menampilkan nilai *mean absolute error*, *mean square error*, dan *explained variance score*.

Berikut adalah diagram perancangan penelitian:



Gambar 3.2 Alur Perancangan Program